

Modelling of word usage frequency dynamics using artificial neural network

Maslennikova Y., Bochkarev V., Voloskov D.

Kazan Federal University, 420008, Kremlevskaya 18, Kazan, Russia

Abstract

In this paper the method for modelling of word usage frequency time series is proposed. An artificial feedforward neural network was used to predict word usage frequencies. The neural network was trained using the maximum likelihood criterion. The Google Books Ngram corpus was used for the analysis. This database provides a large amount of data on frequency of specific word forms for 7 languages. Statistical modelling of word usage frequency time series allows finding optimal fitting and filtering algorithm for subsequent lexicographic analysis and verification of frequency trend models. © Published under licence by IOP Publishing Ltd.

<http://dx.doi.org/10.1088/1742-6596/490/1/012180>
